AD-A053 632   SOUTHERN METHODIST UNIV DALLAS TEX DEPT OF STATISTICS   F/G 12/1
MILKING DATA THROUGH BIASED REGRESSION TECHNIQUES,(U)
1977   R F GUNST, J T WEBSTER   AFOSR-75-2871
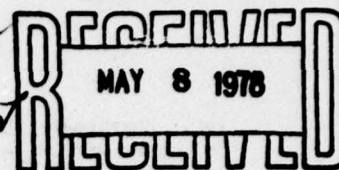UNCLASSIFIED   AFOSR-TR-78-0732   NL

OF
ADA
053632

END
DATE
FILMED

6-78

DDC

MILKING DATA THROUGH BIASED REGRESSION TECHNIQUES

J. T. Webster and R.F. Gunst, Southern Methodist University

The recent interest in biased estimation procedures in multiple linear regression arises from the large variances of the least squares estimators (unbiased) of the regression coefficients when multicollinearities are present. The biased estimation procedures greatly reduce this variance at the cost of some bias. It is the purpose of this paper to look at this bias with reference to the nature of the specific problem being investigated.

For a structure upon which to base this discussion let the model be of the form

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_p X_{ip} + \epsilon_i$$
$$i = 1, 2, \ldots, n$$

or $\underline{Y} = \underline{1}\beta_0 + X\underline{\beta} + \underline{\epsilon} \qquad (1)$

where $\underline{Y}' = [Y_1, Y_2, \ldots, Y_n]$; $\underline{1}$ is an $n \times 1$ vector of ones; β_0 and $\beta' = [\beta_1, \beta_2, \ldots, \beta_p]$ are unknown parameters; $X = \{X_{ij}\}$, an $n \times p$ matrix of known values of rank p with $X'\underline{1} = \phi$ and the diagonal elements of $X'X$ equal to one; and $\underline{\epsilon}' = [\epsilon_1, \epsilon_2, \ldots \epsilon_n]$ with ϵ_i a random variable, $E(\epsilon_i) = 0, E(\epsilon_i^2) = \sigma^2$ and $E(\epsilon_i \epsilon_{i'}) = 0, i \neq i'$.

There are basically two questions addressed in a multiple linear regression analysis: i) The estimation of $\underline{\beta}$, and ii) the estimation of Y, a future value, when $X_1, X_2 \ldots X_p$ are given. The answers are not necessarily the same as will become apparent in our discussion. We shall in essence be looking for a solution to i) but also show that a poor solution to i) may still be a good solution to ii) in a restricted sense.

1. A Decomposition of $\underline{\beta}$

When problems with multicollinearities arise, a tool often used is to express $\underline{\beta}$ as a linear function of the latent vectors of $X'X$ (e.g. Hocking et al. [1976]):

$$\underline{\beta} = \gamma_1\underline{\alpha}_1 + \gamma_2\underline{\alpha}_2 + \ldots + \gamma_p\underline{\alpha}_p \qquad (2)$$

where $\underline{\alpha}_\ell$ is the latent vector of $X'X$ associated with the latent root λ_ℓ. Placing this form of $\underline{\beta}$ into model (1) yields

$$\underline{Y} = \underline{1}\beta_0 + \gamma_1\underline{z}_1 + \gamma_2\underline{z}_2 + \ldots + \gamma_p\underline{z}_p + \underline{\epsilon},$$

where $\underline{z}_\ell = X\underline{\alpha}_\ell$. The $\underline{z}_\ell$ are mutually orthogonal hence the $\gamma_\ell = \underline{z}_\ell'\underline{Y}/\underline{z}_\ell'\underline{z}_\ell$ are uncorrelated, unbiased estimators of γ_ℓ with variance σ^2/λ_ℓ.

A consequence of multicollinearities is some small λ_ℓ (Mason et al. [1975]); hence, some γ_ℓ will have large variances. The least squares estimator of $\underline{\beta}$ will also have some elements with large variances as can be seen from

$$\underline{b} = (X'X)^{-1}X'\underline{Y} = \hat{\gamma}_1\underline{\alpha}_1 + \hat{\gamma}_2\underline{\alpha}_2 + \ldots + \hat{\gamma}_p\underline{\alpha}_p; \qquad (3)$$

giving $Var(b_j) = \sigma^2\left[\dfrac{\alpha_{j1}^2}{\lambda_1} + \dfrac{\alpha_{j2}^2}{\lambda_2} + \ldots + \dfrac{\alpha_{jp}^2}{\lambda_p}\right]. \qquad (4)$

One method of handling this problem is to use a biased estimator which in essence is the same as (3) except replace γ_ℓ with $W_\ell\gamma_\ell$ as the coefficient of $\underline{\alpha}_\ell (0 \leq W_\ell \leq 1)$. Then the corresponding term in (4) is $W_\ell^2 \alpha_{j\ell}^2 / \lambda_\ell$. The idea is to let W_ℓ be small if λ_ℓ is small. Some of the more commonly used biased estimators have W_ℓ of the following forms:

Estimator	W_ℓ
Simple Ridge	$\lambda_\ell/(\lambda_\ell+k)$ $k>0; \ell=1,2,\ldots,p$
Generalized Ridge	$\lambda_\ell/(\lambda_\ell+k_\ell)$ $k_\ell>0; \ell=1,2,\ldots,p$
Principal Components	0 or 1; $\ell=1,2,\ldots,p$
Shrunken	c $0<c<1; \ell=1,2,\ldots,p$

The bias of this estimator of $\underline{\beta}$ would then be

$$\sum_{\ell=1}^{p} (1-W_\ell)\gamma_\ell\underline{\alpha}_\ell.$$

In addition to selecting the W_ℓ so that the variances of the resulting estimators are not extremely large, we wish to choose the W_ℓ so that the bias is not great. Since $\underline{\alpha}_\ell$ is known and not null, in fact $\sum_{j=1}^{p} \alpha_{j\ell}^2 = 1$, information is needed on the γ_ℓ in order to insure that these biases are not so large that the benefits of variances reduction are counteracted.

From (4) we see that we wish to choose W_ℓ small if λ_ℓ is small but from (5) this would inflate the bias unless γ_ℓ is also small. Keep in mind that when λ_ℓ is small $\hat{\gamma}_\ell$ has a large variance, σ^2/λ_ℓ; hence $\hat{\gamma}_\ell$ is of limited value in determining whether γ_ℓ is small. This will be illustrated in the following data.

2. The Mesquite Problem

This data was furnished by R.J. Freund of Texas A & M University. The purpose of the study was in part to determine an estimation equation for the total production of aerial photosynthetic biomass of mesquite from easily measured parameters of the plant. An identification of the variables is given in Table A.1 and the data from 20 trees in Table A.2.

The concept of volume suggests a multiplicative model; hence an analysis of the logarithms of the variables would be appropriate. A linear regression of the raw data yielded approximately the same coefficient of determination, R^2, as a linear regression of the logarithms. For the purpose of this paper we will use the analysis of the raw data since the interpretation of the regression coefficients is less confusing.

The residuals of the least squares fit of the data indicate the second tree to be an out-

lier. Physically this tree also stands out from the others with $X_5 = 9$ primary stems. We discarded this tree as an outlier and continued our analysis using the remaining 19 data points. These had a coefficient of determination of $R^2 = .868$.

Table A.3 contains the standardized $X'X$ matrix. The latent roots of this matrix are $\lambda_1 = 3.33$, $\lambda_2 = .89$, $\lambda_3 = .57$, $\lambda_4 = .12$, $\lambda_5 = .09$, indicating two multicollinearities, although they are not as severe as one often encounters. The variance inflation factors, diagonal of $(X'X)^{-1} = (5.90, 4.58, 6.02, 5.03, 1.37)$, also point out the existence of some multicollinearities and further indicate that plant density, X_5, is not involved in them.

The "total variance" of $\underline{b}$, the sum of the variances of these five estimators, is

$$\sigma^2 \cdot [\text{the sum of the diagonals of } (X'X)^{-1}] = 22.90\sigma^2.$$

This can also be found from

$$\sigma^2 \sum_{\ell=1}^{5} \lambda_\ell^{-1} .$$

From this second form we note that if $W_4 = W_5 = 0$, the resulting biased estimator of $\underline{\beta}$ has total variance

$$\sigma^2 \sum_{\ell=1}^{3} \lambda_\ell^{-1} = 3.46 \; \sigma^2.$$

Thus there is room for a considerable reduction in variance using an estimator of the form

$$\hat{\underline{\beta}} = \hat{\gamma}_1\underline{\alpha}_1 + \hat{\gamma}_2\underline{\alpha}_2 + \hat{\gamma}_3\underline{\alpha}_3 + W_4\hat{\gamma}_4\underline{\alpha}_4 + W_5\hat{\gamma}_5\underline{\alpha}_5. \quad (6)$$

With this in mind, the question then is how much bias would be introduced. The latent vectors $\underline{\alpha}_4$ and $\underline{\alpha}_5$ are given in Table A.4. Note that the fifth element in $\underline{\alpha}_4$ and $\underline{\alpha}_5$ are both relatively small; thus the introduction of small W_4 and W_5 will have little effect upon $\hat{\beta}_5$. Also the variance inflation factor of 1.37 indicates that there is little room for improvement on b_5.

3. An Investigation of the Bias

The bias is affected by both γ_4 and γ_5. We have already noted that $\hat{\gamma}_\ell = \underline{Y}'\underline{z}_\ell / \underline{z}_\ell'\underline{z}_\ell$ is an unbiased estimator of γ_ℓ with variance σ^2/λ_ℓ. Since γ_4 and γ_5 are both quite small, $\hat{\gamma}_4$ and $\hat{\gamma}_5$ both have exceedingly large variances. What this means can be shown graphically.

Due to the orthogonality of the $\underline{z}_\ell$, and since $\underline{z}_\ell' \underline{1} = 0$,

$$\underline{Y}'\underline{z}_\ell = (\underline{Y} - \underline{1}\overline{Y} - \hat{\gamma}_1\underline{z}_1 \cdots - \hat{\gamma}_{\ell-1}\underline{z}_{\ell-1} - \hat{\gamma}_{\ell+1}\underline{z}_{\ell+1} - \cdots - \hat{\gamma}_p\underline{z}_p)'\underline{z}_\ell.$$

The long expression in parentheses on the right is $\underline{Y}$ adjusted for the intercept and all the z's except z_ℓ. Call this partial residual vector $\underline{r}_\ell$. Then $\gamma_\ell = \underline{r}_\ell'\underline{z}_\ell / \underline{z}_\ell'\underline{z}_\ell$ is the least squares slope of r_ℓ versus z_ℓ, a slope that is easily

visualized through a plot of r_ℓ vs. z_ℓ. Figures A1 through A5 show this for the five z's of this example.

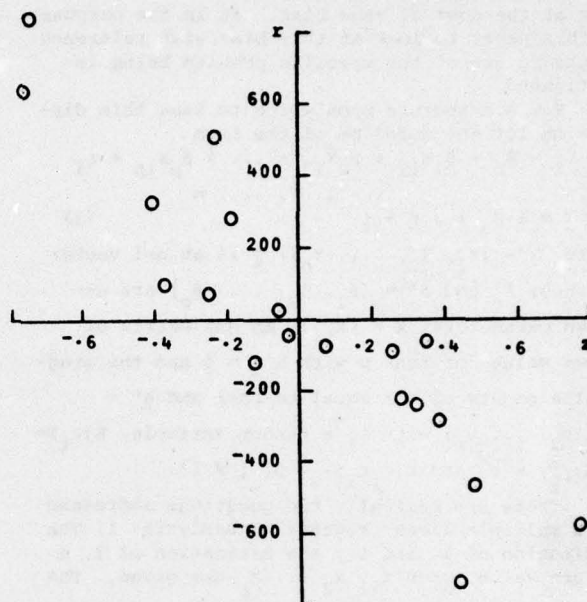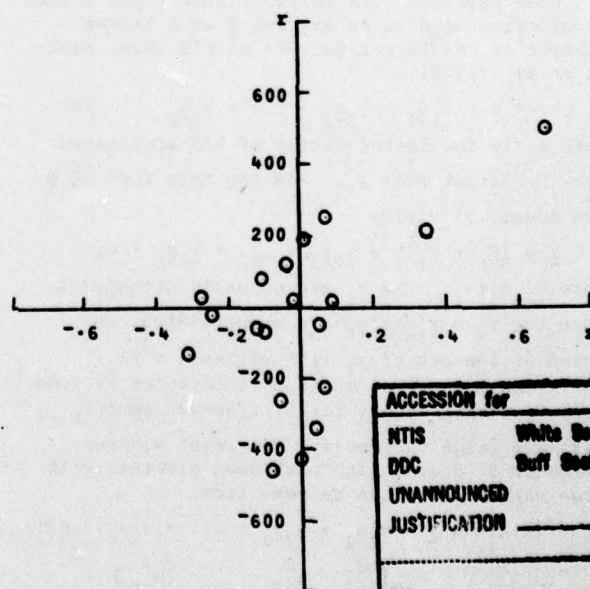#### FIGURE A1. PARTIAL RESIDUAL PLOT OF FIRST PRINCIPAL COMPONENT

FIGURE A2. PARTIAL RESIDUAL PLOT OF SECOND PRINCIPAL COMPONENT

FIGURE A3. PARTIAL RESIDUAL PLOT
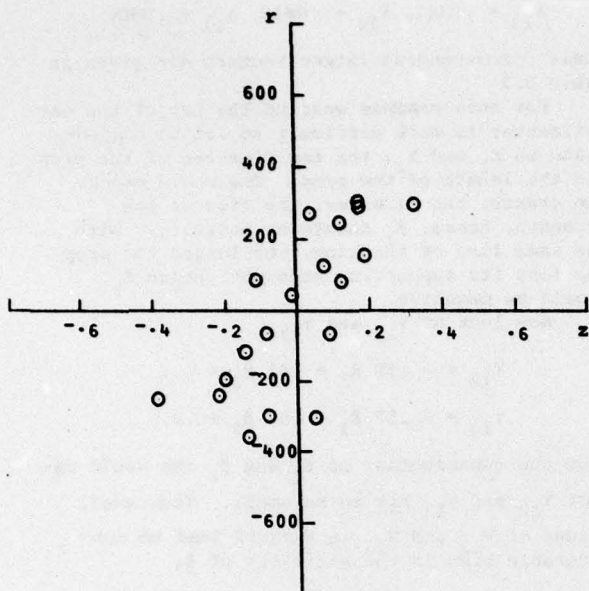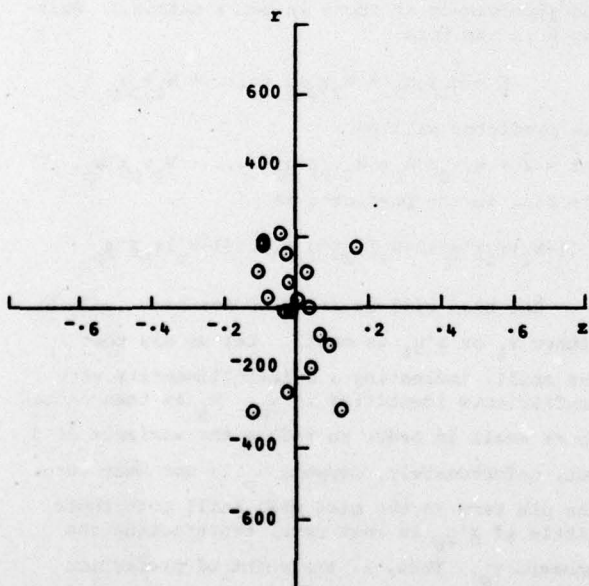## OF THIRD PRINCIPAL COMPONENT

FIGURE A4. PARTIAL RESIDUAL PLOT
OF FOURTH PRINCIPAL COMPONENT

For a moment ignore the source of this data and consider that each of these five plots was brought to you separately and you were asked to find a predictor for each. We believe that for Figures A4 and A5 you would quite politely say that there is not enough information to establish a relationship between the two variables. There may or may not be a relationship but data is needed outside the range of ± .2. For example

FIGURE A5. PARTIAL RESIDUAL PLOT
OF FIFTH PRINCIPAL COMPONENT
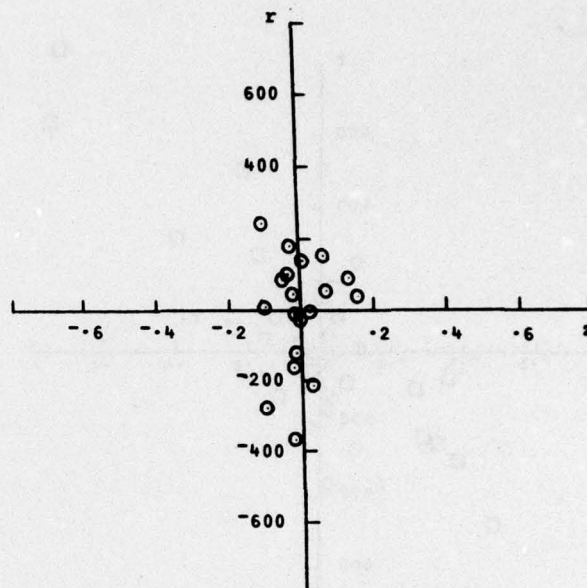
look at Figures A6 and A7. (Figure A6 is the composite of Figures A1 and A5. Figure A7 is the composite of the reflection of Figure A1 and Figure A5). Both of these are compatable with Figure A5 and either one could be the result of data for z_5 outside of ± .2. The crux of the matter is that the available data gives us little or no information on γ_4 and γ_5.

FIG. A6. PARTIAL RESIDUAL PLOT OF FIFTH PRINCIPAL
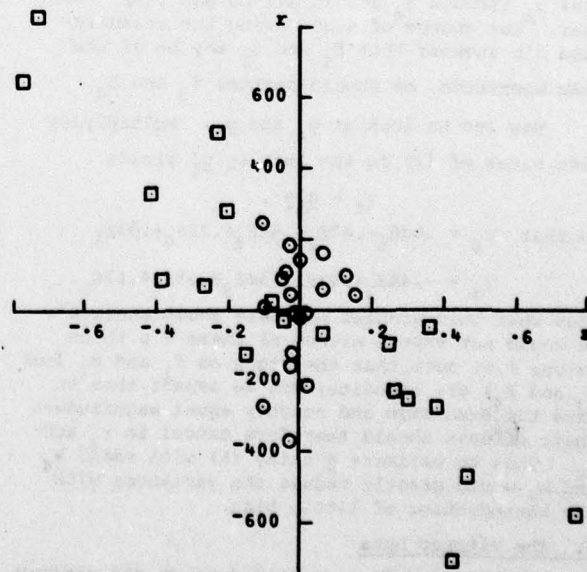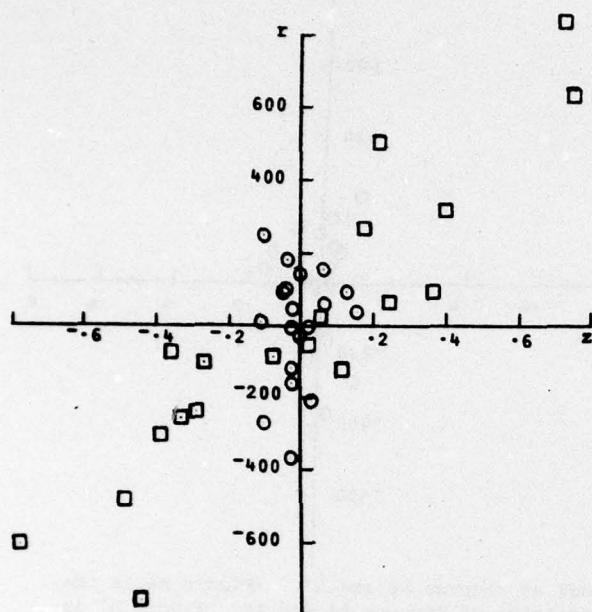## COMPONENT AUGMENTED WITH POINTS FROM FIGURE A1

9

FIG. A7. PARTIAL RESIDUAL PLOT OF FIFTH PRINCIPAL COMPONENT AUGMENTED WITH REFLECTED POINTS FROM FIGURE A1

There is, however, one other available source of information: the problem itself and the experimenter knowledgeable in the area. Because of the simplicity of the measurements in this problem we can perhaps wear his hat, keeping in mind that the variable for prediction is the amount of leaves.

Looking at the nature of X_1 through X_4 we see that all four are size measurements and increasing any one should indicate more leaves. Thus β_1 through β_4 should all be positive. Further, the change of scale using the standardized X's suggest that β_1 and β_2 may be of the same magnitude, as should perhaps β_3 and β_4.

Now let us look at $\underline{\alpha}_4$ and $\underline{\alpha}_5$. Multiplying both sides of (2) on the left by $\underline{\alpha}_\ell'$ yields

$$\gamma_\ell = \underline{\alpha}_\ell' \underline{\beta} ,$$

so that $\gamma_4 = .62\beta_1 - .47\beta_2 - .52\beta_3 + .33\beta_4 + .03\beta_5$

$$\gamma_5 = -.44\beta_1 + .42\beta_2 - .54\beta_3 + .55\beta_4 + .17\beta_5.$$

From what this problem suggests about these β's we would not expect either of these γ's to be large; i.e. note that the signs on β_1 and β_2 (and β_3 and β_4) are opposite, yet we expect them to have the same sign and roughly equal magnitude—their effects should therefore cancel in γ_4 and γ_5. Thus to estimate $\underline{\beta}$ using (6) with small W_4 and W_5 would greatly reduce the variances with the introduction of little bias.

4. The Pitprop Data

Jeffers [1967] presented data on the maximum compressive strength of timbers used in mines. A description of the variables is given in Table B.1 and the standardized X'X matrix in Table B.2. This X'X matrix has three small latent roots

$$\lambda_{13} = .0387, \quad \lambda_{12} = .0415, \quad \lambda_{11} = .0506.$$

Their corresponding latent vectors are given in Table B.3

For this example wearing the hat of the experimenter is more difficult so let us concentrate on X_1 and X_2, the top diameter of the prop and the length of the prop. One would expect the greater the diameter, the greater the strength; hence, β_1 should be positive. With the same line of thinking, the longer the prop the less its supporting strength; hence β_2 should be negative.

Now look at γ_{12} and γ_{13} :

$$\gamma_{12} = -.39 \beta_1 + .41 \beta_2 + \ldots$$

$$\gamma_{13} = -.57 \beta_1 + .58 \beta_2 + \ldots .$$

From the contribution of β_1 and β_2 one would expect γ_{12} and γ_{13} not to be small. Thus small values of W_{12} and W_{13} in $\underline{\hat{\beta}}$ could lead to considerable bias in the estimator of $\underline{\beta}$.

5. The Effect of W_ℓ on the Prediction of Y

Up to this time we have been concerned with the estimation of $\underline{\beta}$. Now let us consider the use of this estimate to predict Y at a point $\underline{X}$. (Note that the X's in this vector are scaled in the same manner as those in our X matrix.) Writing $\underline{\hat{\beta}}$ in the form

$$\underline{\hat{\beta}} = W_1 \hat{\gamma}_1 \underline{\alpha}_1 + W_2 \hat{\gamma}_2 \underline{\alpha}_2 + \ldots + W_p \hat{\gamma}_p \underline{\alpha}_p$$

the predictor will be

$$\hat{Y} = \bar{Y} + W_1 \hat{\gamma}_1 \underline{X}' \underline{\alpha}_1 + W_2 \hat{\gamma}_2 \underline{X}' \underline{\alpha}_2 + \ldots + W_p \hat{\gamma}_p \underline{X}' \underline{\alpha}_p. \quad (7)$$

The bias in the predictor is

$$(1-W_1)\gamma_1 \underline{X}'\underline{\alpha}_1 + (1-W_2)\gamma_2 \underline{X}'\underline{\alpha}_2 + \ldots + (1-W_p)\gamma_p \underline{X}'\underline{\alpha}_p .$$

The bias will be small if for each small W_ℓ either γ_ℓ or $\underline{X}'\underline{\alpha}_\ell$ is small. Let us say that λ_p was small, indicating a multicollinearity with coefficients identified in $\underline{\alpha}_p$. W_p is then chosen to be small in order to reduce the variance of $\underline{\hat{\beta}}$ but, unfortunately, suppose γ_p is not near zero. The pth term in the bias will still contribute little if $\underline{X}'\underline{\alpha}_p$ is near zero, conteracting the nonzero γ_p. Thus, if the point of prediction satisfies this multicollinearity of the original data, a small W_p will induce little bias in the predictor.

To summarize this last paragraph, consider that the W_ℓ are chosen small only when λ_ℓ is small. Then, regardless of the magnitude of the γ_ℓ (within reason), the $\hat{Y}$ of (7) will have little bias if the point of prediction, $\underline{X}$, satisfies the multicollinearities of the original data.

For example a prediction equation from the pitprop data using small W_{12} and W_{13} could be quite satisfactory for props with multicollinearities designated by $\underline{\alpha}_{12}$ and $\underline{\alpha}_{13}$. However interpretation of the $\hat{\beta}_j$ in this equation could be misleading. On the other hand interpretation of the β_j from the mesquite data using small W_4 and W_5 should be informative.

6. Conclusion

In biased linear regression techniques, small W_ℓ are desired in order to reduce the variance of $\underline{\hat{\beta}}$ in the presence of multicollinearities. This will introduce considerable bias in $\underline{\hat{\beta}}$ if the corresponding γ_ℓ are large. The basic point of this paper is that when λ_ℓ is small the data itself generally gives little information on γ_ℓ. This is well illustrated by the graphs of the mesquite data. Some information may be available, however, if γ_ℓ is expressed as a linear function of the elements of $\underline{\beta}$ and the nature of the specific problem is carefully analyzed.

We feel that in practice a complete investigation of the properties of a set of regression data, as suggested above, is sometimes hampered by the lack of computer programs. Robert Pierce, while at SMU, has written a program, REGRESS, which will simultaneously do a least squares, ridge, latent root and principal component analysis as well as furnish a shrunken estimate. The latent roots and vectors of X'X are a portion of the printout. A copy of this program is available upon request.

REFERENCES

Hocking, R.R., Speed, F.J. and Lynn, M.J. (1976), "A class of biased estimators in linear regression," Technometrics, 18, 425-437.

Jeffers, J.N.R. (1967), "Two case studies in the application of principal component analysis," Applied Statistics, 16, 225-236.

Mason, R.L., Webster, J.T. and Gunst, R.F. (1975), "Sources of multicollinearity in regression analysis," Communications in Statistics, 4, 277-292.

TABLE A.1

MESQUITE DATA: RESPONSE AND PREDICTOR VARIABLES

Variable	Description
Y(LEAFWT)	Total Weight (GRAMS) of Photosynthetic Material
X_1(DIAM1)	Canopy Diameter (METERS) Measured Along the Longest Axis of the Tree Parallel to the Ground
X_2(DIAM2)	Canopy Diameter (METERS) Measured Along the Shortest Axis of the Tree Parallel to the Ground
X_3(TOTHT)	Total Height (METERS) of the Mesquite Bush
X_4(CANHT)	Canopy Height (METERS) of the Mesquite Bush
X_5(DENS)	Plant Unit Density (NUMBER OF PRIMARY STEMS/PLANT UNIT)

TABLE A.2

MESQUITE DATA

Bush Number	DIAM1	DIAM2	TOTHT	CANHT	DENS	LEAFWT
1	2.50	2.30	1.70	1.40	5	723.0
2	5.20	4.00	3.00	2.50	9	4052.0
3	2.00	1.60	1.70	1.40	1	345.0
4	1.60	1.60	1.60	1.30	1	330.9
5	1.40	1.00	1.40	1.10	1	163.5
6	3.20	1.90	1.90	1.50	3	1160.0
7	1.90	1.80	1.10	.80	1	386.6
8	2.40	2.40	1.60	1.10	3	693.5
9	2.50	1.80	2.00	1.30	7	674.4
10	2.10	1.50	1.25	.85	1	217.5
11	2.40	2.20	2.00	1.50	2	771.3
12	2.40	1.70	1.30	1.20	2	341.7
13	1.90	1.20	1.45	1.15	2	125.7
14	2.70	2.50	2.20	1.50	3	462.5
15	1.30	1.10	.70	.70	1	64.5
16	2.90	2.70	1.90	1.90	1	850.6
17	2.10	1.00	1.80	1.50	2	226.0
18	4.10	3.80	2.00	1.50	2	1745.1
19	2.80	2.50	2.20	1.50	1	908.0
20	1.27	1.00	.92	.62	1	213.5

TABLE A.3

CORRELATION MATRIX OF PREDICTOR AND RESPONSE VARIABLES, MESQUITE DATA (n = 19)

	X_1	X_2	X_3	X_4	X_5
X_1	1.00	.88	.73	.68	.32
X_2		1.00	.64	.58	.20
X_3			1.00	.88	.39
X_4				1.00	.22
X_5					1.00

TABLE A.4

SMALLEST TWO LATENT ROOTS AND CORRESPONDING LATENT VECTORS OF X'X, MESQUITE DATA (n = 19)

$$\lambda_5 = .0904 \qquad \lambda_4 = .1155$$

VARIABLE	$\underline{\alpha}_5$	$\underline{\alpha}_4$	VIF
X_1	-.444	.624	5.9
X_2	.420	-.472	4.6
X_3	-.544	-.522	6.0
X_4	.551	.339	5.0
X_5	.165	.026	1.4

TABLE B.1

PITPROP DATA: RESPONSE AND PREDICTOR VARIABLES

VARIABLE	DESCRIPTION
Y	Maximum Compressive Strength of Prop
X_1	Top Diameter of the Prop (INCHES)
X_2	Length of Prop (INCHES)
X_3	Moisture Content of Prop (% OF DRY WEIGHT)
X_4	Specific Gravity of the Timber (AT TIME OF TEST)
X_5	Oven-dry Specific Gravity of the Timber
X_6	Number of Annual Rings at Top of Prop
X_7	Number of Annual Rings at Base of Prop
X_8	Maximum Bow (INCHES)
X_9	Distance of the Point of Maximum Bow from Top of Prop (INCHES)
X_{10}	Number of Knot Whorls
X_{11}	Length of Clear Prop from Top of Prop (INCHES)
X_{12}	Average Number of Knots per Whorl
X_{13}	Average Diameter of Knots (INCHES)

TABLE B.2

CORRELATION MATRIX OF PREDICTOR AND RESPONSE VARIABLES, PITPROP DATA

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}
X_1	1.00	.95	.36	.34	-.13	.31	.50	.42	.59	.55	.08	-.02	.13
X_2		1.00	.30	.28	-.12	.29	.50	.42	.65	.57	.08	-.04	.14
X_3			1.00	.88	-.15	.15	-.03	-.05	.13	-.08	.16	.22	.13
X_4				1.00	.22	.38	.17	-.06	.14	-.01	.10	.17	.02
X_5					1.00	.36	.30	.00	-.04	.04	-.09	-.15	-.21
X_6						1.00	.81	.09	.21	.27	-.04	.02	-.33
X_7							1.00	.37	.47	.68	-.11	-.23	-.42
X_8								1.00	.48	.56	.06	-.36	-.20
X_9									1.00	.53	.09	-.13	-.08
X_{10}										1.00	-.32	-.37	-.29
X_{11}											1.00	.03	.01
X_{12}												1.00	.18
X_{13}													1.00
Y	-.42	-.34	-.73	-.54	.25	.12	.11	-.25	-.24	-.10	-.06	-.12	-.15

TABLE B.3

SMALLEST THREE LATENT ROOTS AND CORRESPONDING LATENT VECTORS OF X'X, PITPROP DATA

VARIABLE	$\lambda_{13} = .0387$ $\underline{\alpha}_{13}$	$\lambda_{12} = .0415$ $\underline{\alpha}_{12}$	$\lambda_{11} = .0506$ $\underline{\alpha}_{11}$	VIF
X_1	-.572	-.392	-.005	13
X_2	.582	.411	-.054	14
X_3	.408	-.527	.117	12
X_4	-.383	.585	-.017	12
X_5	.118	-.202	.005	3
X_6	.057	-.080	-.537	7
X_7	.002	.036	.764	12
X_8	.018	.053	.026	2
X_9	-.058	-.054	-.051	2
X_{10}	.004	-.060	-.318	5
X_{11}	-.007	-.005	-.048	2
X_{12}	.004	-.002	.047	1
X_{13}	-.009	-.013	.045	2

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR-78-0732	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MILKING DATA THROUGH BIASED REGRESSION TECHNIQUES		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) R. F. Gunst and J. T. Webster		8. CONTRACT OR GRANT NUMBER(s) AFOSR-75-2871
9. PERFORMING ORGANIZATION NAME AND ADDRESS Southern Methodist University Department of Statistics Dallas, Texas 75275		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research /NM Bldg. 410, Bolling AFB Washington, D.C. 20332		12. REPORT DATE 1977
		13. NUMBER OF PAGES 7
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release: Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Multicollinearity
Regression Analysis
Principal Components

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The recent interest in biased estimation procedures in multiple linear regression arises from the large variances of the least squares estimators (unbiased) of the regression coefficients when multicollinearities are present. The biased estimation procedures greatly reduce this variance at the cost of some bias. It is the purpose of this paper to look at this bias with reference to the nature of the specific problem being investigated.